

AD695808

Professional Paper 29.69

September 1969

# Graph Theory as a Metlanguage of Communicable Knowledge

by

Edward Kingsley, Felix F. Kopstein, and Robert J. Seidel

Presentation at the  
Society for General Systems Research  
Annual Meeting  
Dallas, Texas    December 1968

**HumRRO**

This document has been approved  
for public release and sale;  
its distribution is unlimited.

**HUMAN RESOURCES RESEARCH ORGANIZATION**

|   |   |   |   |   |   |   |   |   |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |     |
|---|---|---|---|---|---|---|---|---|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|-----|
| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 | 29 | 30 | 31 | 32 | 33 | 34 | 35 | 36 | 37 | 38 | 39 | 40 | 41 | 42 | 43 | 44 | 45 | 46 | 47 | 48 | 49 | 50 | 51 | 52 | 53 | 54 | 55 | 56 | 57 | 58 | 59 | 60 | 61 | 62 | 63 | 64 | 65 | 66 | 67 | 68 | 69 | 70 | 71 | 72 | 73 | 74 | 75 | 76 | 77 | 78 | 79 | 80 | 81 | 82 | 83 | 84 | 85 | 86 | 87 | 88 | 89 | 90 | 91 | 92 | 93 | 94 | 95 | 96 | 97 | 98 | 99 | 100 |
|---|---|---|---|---|---|---|---|---|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|-----|

The Human Resources Research Organization (HumRRO) is a nonprofit corporation established in 1969 to conduct research in the field of training and education. It is a continuation of The George Washington University Human Resources Research Office. HumRRO's general purpose is to improve human performance, particularly in organizational settings, through behavioral and social science research, development, and consultation. HumRRO's mission in work performed under contract with the Department of the Army is to conduct research in the fields of training, motivation, and leadership.

The contents of this paper are not to be construed as an official Department of the Army position, unless so designated by other authorized documents.

Published  
September 1969  
by  
HUMAN RESOURCES RESEARCH ORGANIZATION  
300 North Washington Street  
Alexandria, Virginia 22314

### **Prefatory Note**

This paper is based on research performed by the Human Resources Research Organization, Division No. 1 (System Operations) at Alexandria, Virginia, under Work Unit IMPACT, Instructional Model/Prototypes Attainable in Computerized Training.

The paper was presented at the annual meeting of the Society for General Systems Research (SGSR) at Dallas, Texas, in December 1968, by Dr. Kopstein; Edward Kingsley, Felix F. Kopstein, and Robert J. Seidel shared equally in preparation of the paper.

## GRAPH THEORY AS A METALANGUAGE OF COMMUNICABLE KNOWLEDGE

Edward Kingsley, Felix F. Kopstein, and Robert J. Seidel

If one is faced, or pretends to be faced, with the problem of specifying to an automaton-instructor, such as a computer, precisely how it must behave in order to instruct a student in some subject matter or other, a number of issues are brought into sharp focus. Computers, contrary to popular expectations, possess no special magic, and the mere presence of imposing hardware is no guarantee that effective and efficient learning will occur. A meaningful way of describing the role of the computer in computer-administered instruction (CAI) is to characterize it as a tool—possibly an indispensable tool—for gaining and maintaining a high degree of *facilitative* control over the instructional process. If that tool is to be used effectively, it must be told in complete, precise, and totally unambiguous terms what to do. It will not and cannot tolerate obscurity of language in the commands governing its behavior.

### Instruction as the Communication of Knowledge

Any definition of instruction (or teaching) must include the notion of communication of information so as to effect a transfer of knowledge from an instructor to a student. Also, the concept of instruction must not contain anything that is at variance with a rigorous definition of learning. It is certain that the notion of transfer of knowledge, or of a capability for performing, or of a response repertoire, is entirely consistent with definitions of learning generally accepted in scientific psychology (e.g., Kimble, 1, 2).

In formal instruction, as in a classroom or a tutorial situation, at least two aspects of relevant knowledge may be gained by a student (learner). (Nothing will be said about irrelevant knowledge.) First, he will gain information about the instructor and his characteristics

as an instructional agent. Second, it is hoped that he will gain information about the subject matter.

This subject matter can be viewed—for the moment—as information stored by the instructional agent. The central task for the latter is to transfer a reasonably complete copy of his own course-relevant knowledge to the student. If the instructional agent is an automaton, or if one pretends that he is, two problems present themselves immediately. How much of the information stored by the instructional agent is to be communicated? What is, and what is not, part of the instructional message?

In the second place, the capacity of the communication channel is such that a simultaneous transmission of the total message is out of the question. A sequential transmission of parts is the only possible one. How is the automaton to decide what is a part, how many there are, what sequences of these parts are permissible under existing constraints (and what are these constraints), what parts and sequences of parts have been transmitted, which of them have been not only transmitted but received and stored, and when is transmission of the total instructional message complete? Allied with these decisions is the problem of specifying the message units and how the parts or pieces or chunks of knowledge may be encoded into message units to maximize the probability of reception and assimilation.

#### The Subject-Matter Map

Even if one feels repugnance toward the notion of an automaton in the role of the instructional agent (this is a matter for a different debate), a valuable purpose can be served in at least considering the possibility. For if one is to *prescribe* to the automaton how to instruct, it will be necessary to *describe* knowledge largely independent of any particular content. If the knowledge were merely a matter of a listing, as in a dictionary, it might be expedient to represent it as a series of numbers. However, simply storing and describing an encyclopedic stock of information does not sufficiently represent what we really seem to mean by knowledge. Without some

specified subject-matter organization, there exists no basis on which the hypothetical automaton can decide which facts, principles, procedures, and so forth, to present, what interrelationships exist among them, and in what sequence the topics may be best presented. Facts, principles, concepts, and especially the verbal or symbolic representations required to render them communicable, exist in contexts and never in absolute isolation (e.g., Ackoff, 3, pp. 16-17). Depending on the particular context in which they occur, they are linked by one or more relations. As a minimum, such relations are inevitably demanded by the *inescapable* circularity of definitions (Churchman and Ackoff, 4).

What has been said amounts to this: The requirement exists for a metalanguage in which to describe communicable knowledge. A strong candidate for this role is the mathematics of nets and graphs, or more generally, the field of topology. Obviously, there is an advantage in choosing a metalanguage that is unambiguous, and that has had an extensive mathematical development. Moreover, it is linked to (i.e., can be restated in terms of) other forms of mathematics, most readily, to set theory on the one hand and matrix algebra on the other. It amounts to conceiving of knowledge, or at least communicable knowledge, as a space that is structured. This space is represented as a set of points that are interconnected or interconnectable by a set of lines. To illustrate, an arbitrary net is shown in Figure 1. It will be readily apparent that such a representation amounts to a "map" of a knowledge space. Indeed, ordinary road maps are technically graphs.

Whether the choice of nets or graphs as the metalanguage of knowledge is the best possible one, or even a good one, surely will be open to debate for a long time. However, some beginning must be made. At the very least, it promises new ways of

A "Map" of Knowledge Space

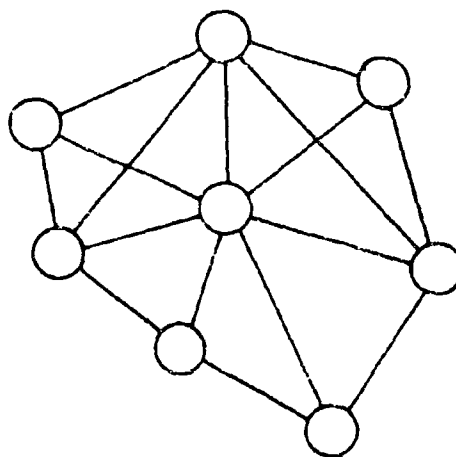


Figure 1

looking at knowledge, and thus may lead to new insights. For example, there are two immediate issues in applying graphs to communicable knowledge qua subject matter. What are the boundaries of the knowledge space and what *exactly* do points and lines represent? Before discussing these issues, however, it may be well to review the nature of graphs and some of their seemingly relevant properties.

### Graphs and Some of Their Properties

Graph theory is concerned with the systematic study of configurations of points and lines joining certain pairs of these points. More formally, consider the structure consisting of a finite set  $V$ , of elements together with a collection,  $U$ , of ordered pairs of elements from  $V$ .

Such a structure is called a *net* by Harary, Norman, and Cartwright (5) and is denoted by  $G = (V, U)$ . The elements of  $V$  are called *points* (*vertices*, *nodes*) and are denoted by  $(v_1, v_2, \dots, v_n)$ . The pairs in  $U$  are called *lines* (*arcs*, *edges*). If  $v_i$  and  $v_j$  are two elements in  $V$ , then the line joining  $v_i$  to  $v_j$  is written as  $(v_i, v_j)$ . Lines are said to be *parallel* if a pair  $(v_i, v_j)$  is repeated in  $U$ . A *loop* is present at a point  $v_i$  if the pair  $(v_i, v_i)$  occurs in  $U$ . A net with no parallel lines is termed a *relation* (Harary, *et al.*, 5). A net with no parallel lines and no loops is called a *directed graph* or *digraph*. (Harary, *et al.*, 5.)

The above definitions are given in terms of set theory. Other representations can be given pictorially and analytically. The pictorial representation of a graph is extremely useful for depicting structural situations. Their map-like nature will be self-evident. In this representation, points are joined by a line or lines associated with them. The relative position of the points is immaterial, as is the shape of the lines joining the points. Lines may cross each other, but such crossovers are not considered as points. A net is shown in Figure 2a, with parallel lines joining  $v_1$  and  $v_2$ , a relation in Figure 2b, with a loop at  $v_2$ , and a digraph in Figure 2c.

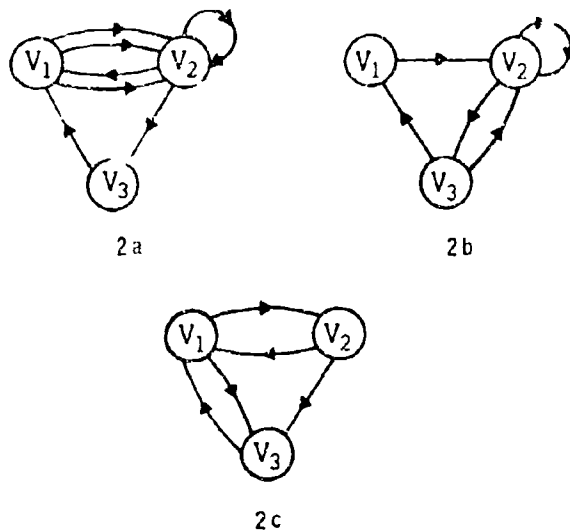
A *partial digraph* of a digraph  $G$  is obtained by using all the points of  $G$  and deleting at least one line. A *subdigraph* of a digraph  $G$  is

obtained by deleting one or more points from  $G$  and their adjacent lines. Possibly partial graphs and subgraphs might be interpreted as regional or not fully developed maps of knowledge.

Figure 3a is one of the partial graphs of the digraph shown in Figure 2c and Figure 3b is one of its subgraphs.

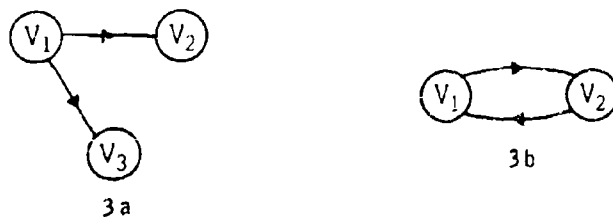
A *path* is a sequence of lines and points in a digraph so that the terminal point of each line is

**Nets With Parallel Lines, Loops, and a Digraph**



**Figure 2**

**Partial Graph and Subgraph of Digraph**



**Figure 3**

coincident with the initial point of the succeeding lines. The *length* of a path is the number of lines in the sequence. In the digraph shown in Figure 4a, there are two paths from  $v_1$  to  $v_3$ , one path has length 1 and the other path, length 2. Possibly paths and their lengths might describe sequences of concepts or topics (points) and the instructional steps separating them (lines).

Two points  $u$  and  $v$  are *adjacent* if they are joined by at least one line. A digraph is *symmetric* if two adjacent points  $v_i$  and  $v_j$  are always joined by two oppositely directed lines. A symmetric digraph is shown in Figure 4b.



### Paths in a Digraph, Symmetric Digraph, and Complete Digraph

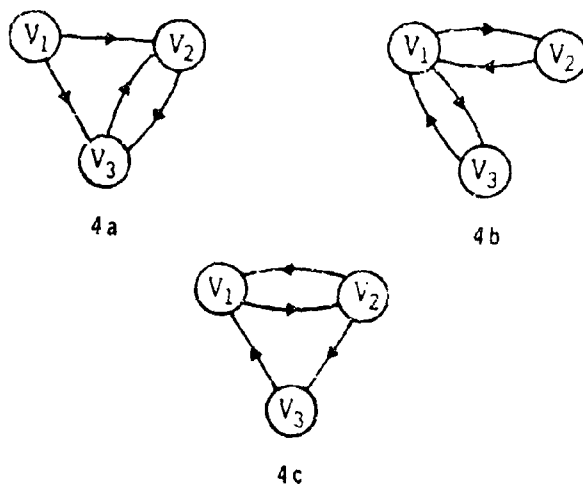


Figure 4

Two topics in a course of instruction might be considered to represent adjacency if one immediately followed the other along a path leading to maximal end-of-course proficiency. In like manner, symmetry would be represented if either topic could be taught before the other with equal effectiveness.

A digraph is *complete* if every pair of points is joined by at least one line oriented in one of the two

directions. A complete graph is illustrated in Figure 4c.

If a digraph is both symmetric and complete—it is called a complete symmetric digraph. The complete symmetric digraph of order 3 is shown in Figure 5.

A digraph is *strongly connected* if, for every two points  $v_i$  and  $v_j$ , there exists a path directed from  $v_i$  to  $v_j$ , as shown in Figure 6. Completeness, symmetry, and strength might be viewed as characteristics that distinguish inherently highly structured subject matter, such as mathematics, from inherently more unstructured subject matter such as history.

### Complete Symmetric Digraph

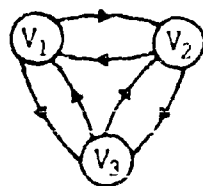


Figure 5

### Strongly Connected Digraph

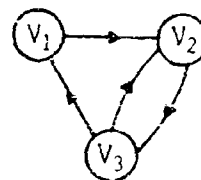


Figure 6

A method for simplifying a digraph that is useful for obtaining insight into its structural properties will be described. Let  $G = (V, U)$  be a digraph and partition its point set  $V$  into  $m$  subsets  $\Pi_1, \Pi_2, \dots, \Pi_m$  so that

$$\bigcup \Pi_i = V$$

$$\Pi_i^{lm} \cap \Pi_j = \phi, i = j; i, j = 1, 2, \dots, m.$$

Then consider the  $m$  subsets of the partition as the points of a new digraph. Each point,  $\Pi_i (i = 1, 2, \dots, m)$ , being labeled by the set of elements of  $G$  that form  $\Pi_i$ . Finally, a line exists from point  $\Pi_i$  to point  $\Pi_j$ , if, and only if, there exists at least one line in  $G$  from a point of  $\Pi_i$  to a point of  $\Pi_j$ . The digraph,  $G^* = (V^*, U^*)$ , thus formed is called the *condensation* of  $G$  with respect to the partition. The point set of a digraph can be partitioned in many different ways. For this reason, there exists a variety of condensations of a digraph. This is illustrated for the digraph  $G$  in Figure 7. Two different condensations,  $G_1^*$  and  $G_2^*$ , are shown for  $G$ . The partitions of  $\{v_1, v_2, v_3, v_4, v_5\}$ , the points of  $G$  that form the points of  $G_1^*$  and  $G_2^*$  are:

$$\Pi_1 = \{v_1, v_2, v_3\}, \Pi_2 = \{v_4, v_5\} \text{ for } G_1^*,$$

$$\Pi_1^1 = \{v_1, v_2\}, \Pi_2^1 = \{v_3\}, \Pi_3^1 = \{v_4, v_5\} \text{ for } G_2^*.$$

Complex digraphs can be made progressively simple by repeated use of the condensation method, a technique that is particularly useful in

### Condensations of a Digraph

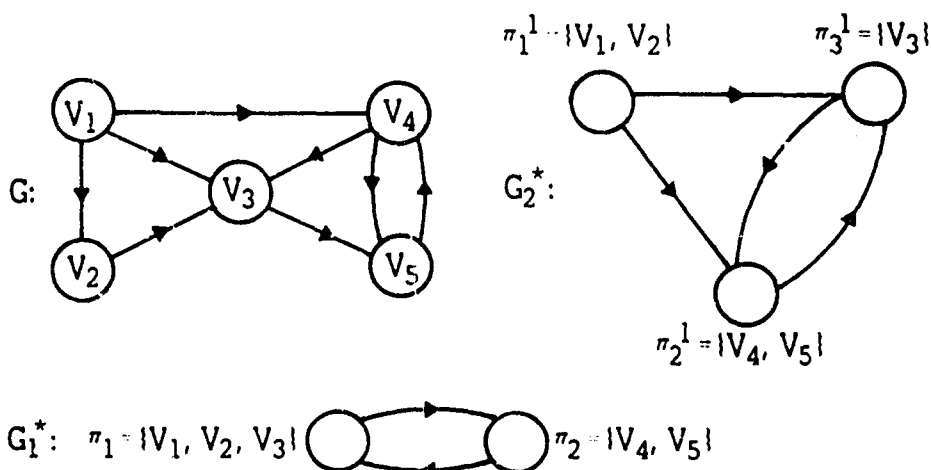


Figure 7

studying hierarchical structures. A possible practical application will be discussed briefly later. Figure 8 illustrates two successive condensations. First, the basic digraph  $G$  is condensed to  $G_1^*$  and

### Two Successive Condensed Digraphs

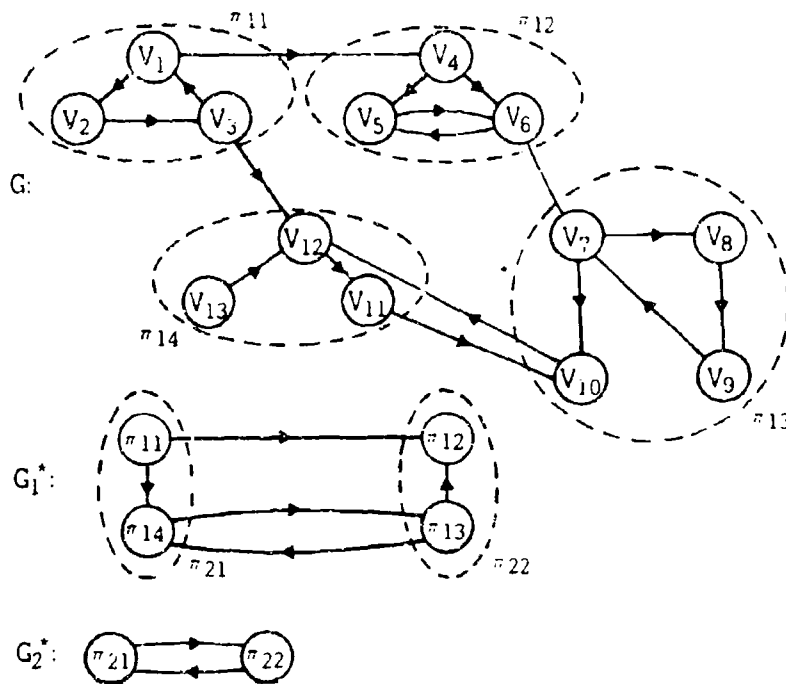


Figure 8

then  $G_1^*$  is condensed next to  $G_2^*$ . The points of  $G_1^*$  are sets of points of  $G$  and the points of  $G_2^*$  are sets of sets of points of  $G$ . The first condensed digraph  $G_1^*$  is formed by the partition:

$$\pi_{11} = \{v_1, v_2, v_3\},$$

$$\pi_{12} = \{v_4, v_5, v_6\},$$

$$\pi_{13} = \{v_7, v_8, v_9, v_{10}\},$$

$$\pi_{14} = \{v_{11}, v_{12}, v_{13}\}.$$

The second condensed digraph,  $G_2^*$ , is formed by partitioning the points of  $G_1^*$ ,  $\{\pi_{11}, \pi_{12}, \pi_{13}, \pi_{14}\}$ , by

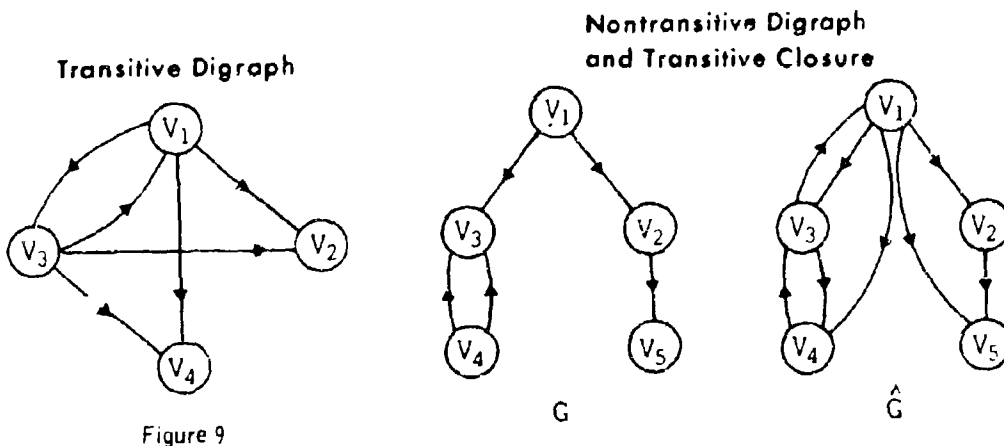
$$\pi_{21} = \{\pi_{11}, \pi_{14}\} = \{v_1, v_2, v_3\}, \{v_{11}, v_{12}, v_{13}\},$$

$$\pi_{22} = \{\pi_{12}, \pi_{13}\} = \{v_4, v_5, v_6\}, \{v_7, v_8, v_9, v_{10}\}.$$

It will be reasonably self-evident that the condensation procedure is an excellent analog to abstraction. With each successive partitioning and condensation, any point in the condensed graph is likely to become more inclusive. At the same time, there is a progressive loss of information about the subgraph(s) that have been condensed. This would seem to suggest that a concept at a given level of abstraction should be presented only after the component subconcepts and their interrelations have been taught.

A digraph  $G$  is transitive if, for every three points  $v_i, v_j, v_k$  in  $G$ , it contains the line  $v_i v_k$  whenever it contains both of the lines  $v_i v_j$  and  $v_j v_k$ . In other words, if there is a path of length 2 from point  $v_i$  to point  $v_k$  then there exists a line from  $v_i$  to  $v_k$ . The digraph shown in Figure 9 is a transitive digraph.

If a digraph  $G = (V, U)$  is not transitive, a new transitive digraph can be obtained from it by constructing the minimal transitive digraph containing  $G$  and leaving the same vertices as  $G$ . The digraph,  $\hat{G} = (V, \hat{U})$  thus formed is called the *transitive closure* of  $G$ . Figure 10 shows a



nontransitive digraph  $G$  and its transitive closure  $\hat{G}$ . The digraph  $\hat{G}$  has the same point set as  $G$  but a different set of arcs.

Provisionally the transitive closure procedures might be seen to have application to such issues as a student developing new "insights" into interrelations among concepts or principles within a subject matter. This will be further explored in the context of relative knowledge space.

If a digraph contains the line  $v_i v_j$  but not the line  $v_j v_i$ , the digraph is said to be *asymmetric*. In other words, if two points  $v_i$  and  $v_j$  are connected by a line, there is only one line, either from  $v_i$  to  $v_j$  or from  $v_j$  to  $v_i$ .

If a digraph is both transitive and asymmetric, it is termed the digraph of a *partial order*, as shown in Figure 11.

Digraph of a Partial Order

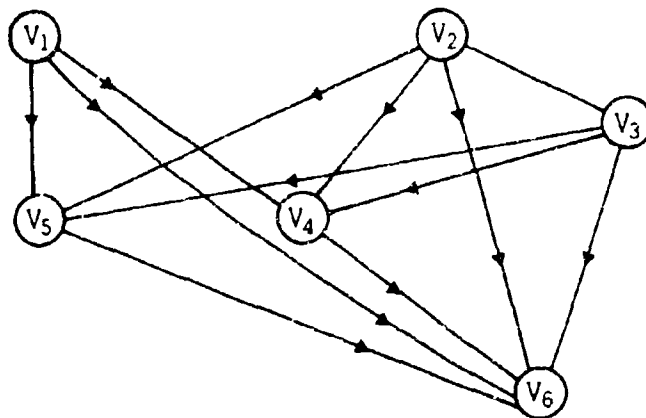


Figure 11

It is often useful to be able to arrange the points of a partially ordered digraph in a linear sequence without destroying the partial ordering. That is, the points in the digraph are rearranged so that all the arrows go from top to bottom, resulting in a *linearly sorted* digraph. The digraph has not changed by this process. A particular digraph of a partial order may have several linearly sorted digraphs, as illustrated in Figure 12 for the partially ordered digraph  $G$ . Three of the many linearly sorted digraphs  $S_1, S_2$ , and  $S_3$  are given for  $G$ .

### Linearly Sorted Digraph

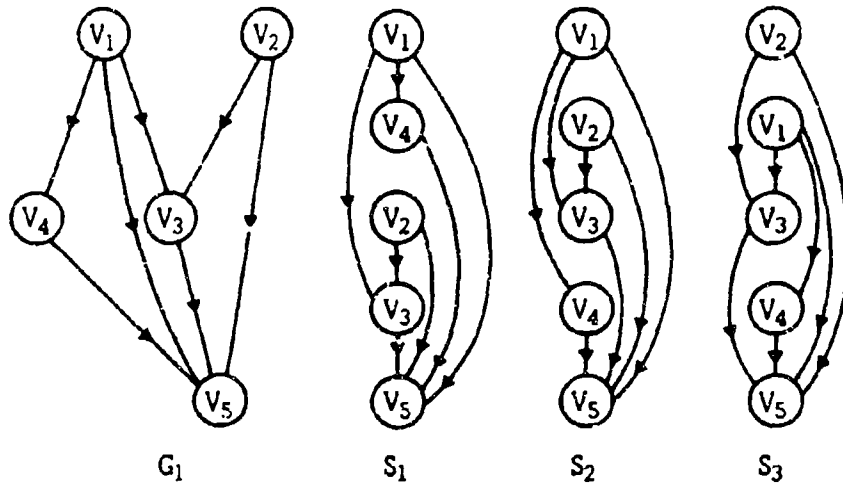


Figure 12

Linear sorts will be readily identified with specific "routes" through a graph. Where a graph represents, for example, an area of knowledge qua subject matter in a course, the totality of different linear sorts will be equal to the number of different ways in which that subject matter can be presented. The choice among the available routes (linear sorts) might depend on the personal characteristics of an individual student, his within-course history (response patterns), his motivations or preferences, or the preferences or constraints imposed on the instructional agent (teacher, computer program).

### Relative Knowledge Spaces and Subject-Matter Graphs

Earlier the question of boundaries on a knowledge space was raised. This problem applies to a total knowledge space, encompassing all communicable human knowledge, as well as to any subject matter in a particular course of instruction.

Operating premises for a pragmatic, conceptual approach to content-independent descriptions of knowledge need to take the following or similar forms: First, any given subject matter is an *arbitrarily delimited portion* of the totality of human knowledge. Second, just as all human knowledge is incomplete and constantly increasing, so is

knowledge of the specific subject matter at hand. Increase in knowledge takes place when new facts are discovered and new concepts are established, or when new relationships among existing facts and/or concepts are discovered. For example, in biochemistry recently established facts about the molecular structure of certain substances, DNA and RNA, are being related to a vast number of other facts and concepts on heredity, evolution of species, cancer, the common cold, and so on. In time, we will become aware of other relationships of the structure of DNA with known and yet-to-be discovered facts in many disciplines.

The DNA example also illustrates that divisions among disciplines and subject-matter areas are arbitrary. Good arguments can be presented for assigning the molecular structure of DNA to the subject matter of physics, of chemistry, of physiology, of genetics and biology at large, or even—perhaps, for example, in the process of remembering *qua* information storage and retrieval—to behavioral science.

We must distinguish among (a) the total knowledge that is possessed by mankind so far, (b) arbitrarily delimited subject-matter areas within the totality of existing human knowledge, and (c) the knowledge about a given subject-matter area possessed by a specific individual or group of individuals (e.g., an instructor or instructional staff). The knowledge or information about a subject matter possessed by even the most knowledgeable instructor is, in fact, a finite and even quite limited structure. With reference to a *delimited* knowledge structure (e.g., the arbitrary requirements for a course), success in instruction or the communication of knowledge is shown by an inability (in a "loose" sense) to distinguish between the instructor and the student. That is, the student should have learned all that the instructor knows about the subject matter relevant to the course requirements, and both of them should be able to answer questions or perform tasks pertinent to the requirements of the course equally well.

Although in this sense for the instructor and student to be indistinguishable is an ultimate ideal of instruction, even a close approximation is not readily attained. Especially in the early stages of instruction, only a small portion of the instructor's knowledge structure will have been transferred to the knowledge space of the student. The

structure that will have been transferred will be limited in the number of its points (i.e., only part of the facts or concepts) as well as in its strength (i.e., awareness of interrelations).

The distinctions among the several versions of a subject-matter structure can be viewed as a sequence of partial graphs as in Figure 13.

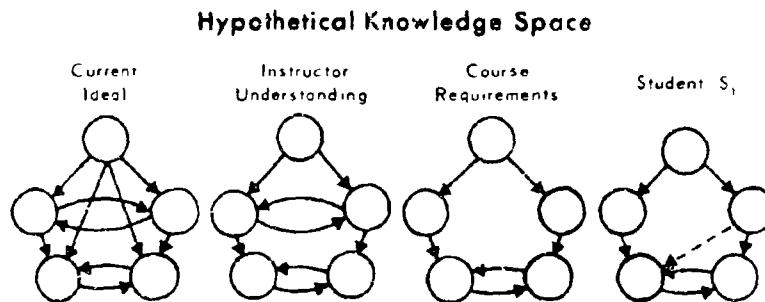


Figure 13

The points represent the concepts of the knowledge space, and the oriented lines, the relations among the concepts. The "Current Ideal" label reflects the fact that knowledge of any given subject matter is constantly "growing." As for the "Instructor Understanding"—this instructor does not have *all* the knowledge about the subject matter that can be possessed currently, but he has enough to serve the requirements of the course; and, also, although it is not required to convey the course properly to the students, he has additional relational awareness. Finally, Student  $S_1$  has not only successfully completed the course requirements, but has independently discovered a *new* relation. (He must be a bright student, since it is a relationship not known to his instructor, and not even existing within the "Current Ideal" structure of the subject matter.)

This simple example is intended to illustrate the necessity for considering *relative* states of knowledge about given subject matter for instructor and student. It accounts for potential individual differences among students. Finally, implicit in it is the notion of generalization—a transfer of potential capability that will engender proficiency in tasks not yet encountered. This capability is important



because the course requirements, as indicated, provide only a subset of the potential relations.

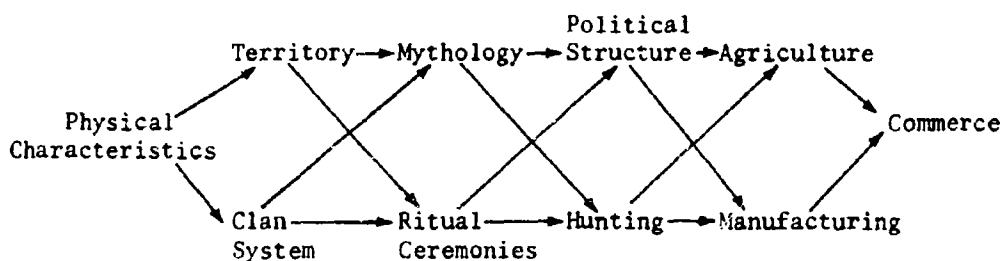
### Graphs and Empirical Reality

The question that remains unanswered concerns the empirical counterparts of points and lines. To be useful, maps need not be simply scaled-down versions of terrain, but they must retain an isomorphism with some aspects of it. For example, road maps can eliminate all cues as to topography and still guide a motorist from town A to town Z. Of course, on arrival at Z the point that represented Z on the regional map will lose its usefulness to the hypothetical motorist who will then need an enlarged city map. This familiar characteristic of road maps corresponds to the condensation and de-condensation procedures in graphs discussed earlier. Also, the representation by points and lines on a regional map is *not* absolutely identical with what is represented by them in the city map. Moreover, this would be true even if the enlarged city representation were an integral part of the regional map.

It appears that maps can incorporate a hierarchy of abstractions without demanding complete consistency from level to level. This realization guided one approach to answering the original question. Kopstein and Hanrieder (6) prepared a pseudo-anthropological account of the culture of a fictitious tribe—the Gruanda. Since it was proposed to use this account as the instructional material in experimental studies of learning as a function of certain structural properties of subject matter, it was necessary to control these structural characteristics. Control could be exercised only by inventing the subject matter to fit the desired structure.

Ten *topics* were arbitrarily chosen, dealing with various aspects of the Gruanda tribe such as their territory, clan system, mythology, rituals, hunting, and agriculture. Each topic constituted a paragraph of about 200 words, and each topical paragraph could stand alone and independent, or could be related to one or more of the other topics. In the latter case, the paragraph contained some sentences that constituted a cross-reference to one or more other topics—for example, the

topic of rituals might contain references to hunting and/or to political structure. The resulting maximal structure follows:



Kopstein and Hanrieder (6) identified topics or paragraphs with points and sentences involving cross-references to other paragraphs as lines. They were clearly aware that this isomorphism reflected only one, relatively high level of abstraction, and thus referred to it as a *macro-structure*.

By implication, this acknowledged underlying *micro-structures*--a recognition that entered into the experimental procedures within which this artificial subject matter was used. Presumably, topical paragraphs are supra-organizations of syntactical and grammatical structures (or, rather, their corresponding behaviors) that are in turn supra-organizations of lexical structures, and so forth. Of course, the reference to the nature of the infra-organization is purely conjectural and intended only as an illustration. Any serious propositions concerning hierarchical levels of structuring would demand a formally consistent taxonomic array and criteria for assigning given structures (graphs) to a particular level. How far the infra-structures may extend downward is an obvious question that need not be settled here.

While experimental data from a subsequent larger study by Kopstein and Seidel failed to support hypotheses which essentially predicted that with higher degrees of structuring, learning or, rather, recall of information about the Gruanda would improve, the validity of these hypotheses remains a moot point. Failure to obtain support may be due to the fact that the operational coordinating definitions were faulty. It is also possible that due to the low complexity and restricted scope of the "Gruanda Material," participating subjects imposed their own structure--perhaps micro-structure. Thus, resolution of the

issues must await further experimental attempts requiring the mastery of larger amounts of more complex subject matter.

A quite different approach to achieving congruence between formal and empirical structures has been proposed by Regnier and de Montmollin (7):

All material to be taught can be characterized in precise fashion by a terminal behavior that is specified *a priori* and then partitioned, more or less finely, into elements that we will call units of knowledge which constitute an enumerable set . . . we will define the unit of knowledge, whatever the level on which it may be situated in the hierarchy of knowledge, as the putting into relation of at least two terms and therefore being able to give, at least, an 'intelligent' response, i.e., a response characterized by the fact that it puts into play long and indirect circuits and, therefore, is indistinguishable from reflex actions, innate or conditioned, which follow immediately upon the presentation of a stimulus.

Even allowing for the ambiguities deriving from translation of French into English, it would seem that Regnier and de Montmollin assert some questionable mathematical propositions. For example, they state that a relation is "non-reflexive" and "anti-symmetric," when subsequent discussion suggests that it is "irreflexive" and "asymmetric" (the latter designations are used below). Some definitions (e.g., "simply connected") are not clear. They do propose a unit of knowledge that has a striking resemblance to the familiar S-R in psychology. S and R constitute the elements in the set with a binary relation on the set so defined that, when it is empirically demonstrated, it simultaneously verifies the existence of the unit of knowledge. Note that neither S nor R involves anything that can be pointed to in the physical world except in terms of S being the necessary and sufficient condition for R.<sup>1</sup> Also, S and R are clearly conceived as being overt and observable.

It has been pointed out that graphs are relations defined on sets of points. Regnier and de Montmollin define a relation  $\mathcal{L}$  on the set of units of knowledge:

It is necessary to have acquired the unit of knowledge  $x_i$  (i.e., to have correctly responded to the question or to

---

<sup>1</sup>Scandura (8) has proposed a set-function language (SFL) that would seem to be pertinent here. In SFL, the basic concern is not with S and R per se, but with the rules relating these sets.

the problem which  $x_i$  represents) in order to acquire the unit of knowledge  $x_i$  (i.e., in order to respond correctly to the question or to the problem which  $x_i$  represents).

These authors then note that the relation  $\mathcal{L}$  is non-reflexive, asymmetric, and transitive, and is therefore a relation of partial order. Since their problem is to try to determine *a priori* the linear order in which the instructional agent's knowledge structure should be "copied" and transmitted to the student, they are at a loss when  $\mathcal{L}$  produces a partial order and does not suffice to determine a Hamiltonian path through the resultant graph—that is, a path traversing every point once.

Thus a second relation  $P$  is defined as:

The acquisition of the unit of knowledge  $x_i$  effected immediately before the acquisition of a unit of knowledge  $x_j$  facilitates the latter.

Relation  $P$  is transitive and neither symmetric nor asymmetric. Therefore, it is a relation of quasi-order that includes  $\mathcal{L}$ , although  $\mathcal{L}$  takes priority over  $P$ .

Although their experimentation incorporates some obscure aspects (e.g., the precise differentiation of experimental study materials) and their measure of the structuring of their materials suffers from mathematical defects noted above, Regnier and de Montmollin's experiment is another initial attempt to study a largely unexplored and potentially fruitful domain. Their preliminary finding was that "the performance on the final test . . . is proportionally inverse to the value of  $\mathcal{O}$ ."

$\mathcal{O}$  is the symbol for their measure of structuring. They state that this outcome was "exactly contrary to our hypothesis."

A third approach to representing knowledge in terms of graphs has been outlined by V.N. Pushkin (9). It would seem that Pushkin's proposals derive from the attempt to describe creative problem solving by humans as, for example, in certain conditions within chess games. Thus, the instructional agent is neither human, nor a simulation, but an objective problem situation faced by a person who is in a role analogous to that of the student in the previously mentioned study. The thrust of Pushkin's argument is that the position of Newell, Shaw, and Simon (10) that equates heuristic computer programs with human behavior is not

tenable. First, there is a distinction to be made between an object or situation with given properties existing in the real world, and the representation of that object or situation within the problem solver. Second, whereas heuristic programs selectively eliminate unproductive moves toward a goal state in a series of sequential steps, humans pre-establish a set of solution-paths qua representations—a solution map—for potential execution. The latter assertion is supported by eye-movement data obtained from chess players. It is the set of problem solutions qua paths to the goal that are represented or representable as graphs.

In Pushkin's approach, the points of a graph  $G$  would seem to be the possible observable states that the problem object or situation can assume, or, at least, a subset thereof that are relevant to the problem. One point represents the initial, the given state, and another represents the final, or desired state. Given existing constraints, there is some finite number of intermediate states that the situation can assume. Lines would seem to be the relational operators that produce the transformation of any given state into an adjacent one. At any time, there is a graph  $G$  that represents *all* possible valid solution-paths from the initial to the final state. Pushkin seems to argue that out of his representations the human problem-solver forms a problem model consisting of a set of solution-paths that are a subgraph or a partial graph or a partial subgraph  $g$  of graph  $G$ .

Mention must also be made of Tesler, Enea, and Colby's (11) attempt to represent belief systems as directed graphs. Whether belief systems such as a psychiatric patient's view of the world are equivalent to objective knowledge is an open question. Certainly it might be viewed this way, if a psychiatrist sought to describe his patient's belief system to third parties. Tesler *et al.* identify "concepts" with points and "simple relationships" with lines. Concepts can be *sets*, *individuals*, and *propositions*. Directed lines connect concepts whenever any relationship exists. However, lines are more specifically characterized through associated symbols representing the "circumstance" of that line, the type of relation, attitude toward the relation, and so forth. The approach of Tesler *et al.* suffers from certain defects in

its axiomatization, but constitutes still another application of graph theory to represent knowledge.

### Review

This presentation should be viewed as an argument for the *promise* of graph theory as a metalanguage of communicable knowledge—in no sense as complete or final. The problem of identifying the points and lines of graphs with some unambiguous empirical reality as, for example, meaningful instructional subject matter, remains as the central issue. This is not to be confused with the different although related problem of the *levels* of micro-structures addressed earlier. If we can ever hope to obtain a consistency in application to meaningful material, we must be able to develop a rule or set of rules for defining conceptual units within the instructional content.

The Kopstein and Hanrieder study (6), and recent studies by Kopstein and Seidel provide good illustrations of this problem. Predicted differences due to varied graph structure were not obtained. The question arises, why not? Since the graph structures observed all relevant axioms and theorems, the answer must be in the rules coordinating formal representations with empirical structures. In any empirical application, a legitimate question to be raised is: How does the experimenter decide whether one, two, or  $n$  sentences are required for adequate cross-referencing? Secondly, one could also ask: What position in the paragraph would create the best salience for the cross-reference material to transmit the desired message, that is, what organization will permit the best Gestalt or figure-ground representation? Finally, is there a "natural" or consensual ordering across people appropriate to the topical paragraphs used? In the Kopstein and Hanrieder experimental material (6), a rough, after-the-fact assessment seemed to support the existence of such an ordering. In fact, these conditions were corrected in the subsequent studies, and the preliminary results seem to indicate verification of our speculations.

This point is raised to emphasize the fact that one cannot simply develop *a priori* a logically tight theory and expect it to be useful

without establishing equally tight rules of representation in the coordinating definitions of the theory. It is difficult to accomplish this with simple, artificial verbal learning tasks, that is, isolated words or phrases, or short paragraphs. It is particularly difficult—and yet even more important—to achieve with higher order conceptual groupings such as those existing in most classroom instructional materials. The dimensions that must be accounted for in the points, or elementary conceptual units in the latter instance, may well require the application of  $n$ -dimensional topological theory rather than simple graph structures. For example, the construction of instructional content itself establishes contextual linkages within the material—these linkages potentially may require one or more additional dimensions for adequately specifying what is represented by a point. We are currently pursuing this and other possibilities in characterizing the subject-matter structure of a computer administered course we are developing to teach computer programming in COBOL, a higher order computer language.

# LITERATURE CITED

1. Kimble, G.A. *Hilgard and Marquis' Conditioning and Learning*, Appleton-Century-Crofts, Inc., New York, 1961, p. 481.
2. Kimble, G.A. "Categories of Learning and the Problem of Definition," in *Categories of Human Learning*, A.W. Melton (Ed.), Academic Press, New York, 1964, p. 33.
3. Ackoff, Russell L. *Scientific Method: Optimizing Applied Research Decisions*, John Wiley and Sons, Inc. New York, 1962.
4. Churchman, C. West, and Ackoff, R.L. *Psychologistics*, University of Pennsylvania Faculty Research Fund, Philadelphia, Pa., 1947 (Mimeographed).
5. Harary, Frank, Norman, Robert Z., and Cartwright, Dorwin. *Structural Models: An Introduction to the Theory of Directed Graphs*, John Wiley and Sons, Inc., New York, 1965.
6. Kopstein, Felix F., and Hanrieder, Barbara D. *The Macro-structure of Subject Matter as a Factor in Instruction*, Research Memorandum, RM-66-25, Educational Testing Service, Princeton, N.J. November 1966.
7. Regnier, Jacques, and de Montmollin, Maurice. *Reconnaissance de l'organisation, recherche de l'ordonnement des elements et choix du mode d'enseignement de la matiere*, Societe d'Economie et de Mathematique Appliquees, Paris, 1968 (Mimeographed).
8. Scandura, Joseph. "New Directions for Theory and Research on Rule Learning: I. A Set-Function Language," *Acta Psychologica*, vol. 28, 1968, North Holland Publishing Co.
9. Pushkin, V.N. "Psikhologiya myshleniya i printsipy evristicheskogo programmirovaniya," *Voprosy Psikhologii*, vol. 13(6), 1967, pp. 101-117.
10. Newell, A., Shaw, J.C., and Simon, H.A. "A Variety of Intelligent Learning in a General Problem Solver," in *Self-Organizing Systems*, M. Yovitts and S. Cameron (Eds.), Pergamon Press, New York, 1968.
11. Tesler, Lawrence, Enea, Horace, and Colby, Kenneth Mark. *Directed Graph Representation for Computer Simulation of Belief Systems*, Department of Computer Science, Stanford University, Stanford, Calif., 1967.



Unclassified

Security Classification

| DOCUMENT CONTROL DATA - R & D   |  |   |
|---|--|---|
| (Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)  |  |   |
| 1. ORIGINATING ACTIVITY (Corporate author)<br>Human Resources Research Organization<br>300 North Washington Street<br>Alexandria, Virginia 22314  |  | 2a. REPORT SECURITY CLASSIFICATION<br>Unclassified  |
|   |  | 2b. GROUP   |
| 3. REPORT TITLE<br>GRAPH THEORY AS A METALANGUAGE<br>OF COMMUNICABLE KNOWLEDGE  |  |   |
| 4. DESCRIPTIVE NOTES (Type of report and inclusive dates)<br>Professional Paper   |  |   |
| 5. AUTHOR(S) (First name, middle initial, last name)<br>Edward Kingsley, Felix F. Kopstein,<br>and Robert J. Seidel   |  |   |
| 6. REPORT DATE<br>September 1969  | 7a. TOTAL NO. OF PAGES<br>24   | 7b. NO. OF REFS<br>11   |
| 8a. CONTRACT OR GRANT NO.<br>DAHC 19-69-C-0018  | 9a. ORIGINATOR'S REPORT NUMBER(S)<br>Professional Paper 29-69                |   |
| b. PROJECT NO.<br>2Q063101D734  |  |   |
| c.  | 9b. OTHER REPORT NO.(S) (Any other numbers that may be assigned this report) |   |
| d.  |  |   |
| 10. DISTRIBUTION STATEMENT<br>This document has been approved for public release and sale;<br>its distribution is unlimited.  |  |   |
| 11. SUPPLEMENTARY NOTES<br>Presentation at Society for General<br>Systems Research annual meeting,<br>Dallas, Texas, December 1968  |  | 12. SPONSORING MILITARY ACTIVITY<br>Office, Chief of Research and Development<br>Department of the Army<br>Washington, D.C. 20310 |
| 13. ABSTRACT<br>The attempts to devise and develop complete (i.e., integrated) computer-administered instruction (CAI) systems have shown the need for an objective, rigorous, and subject-matter independent means for describing the organization of instructional content. Similar approaches to the problem, adopted independently in the U.S. and in France, involve the establishment of a set of subject-matter terms, concepts, topics, or other "units," and the subsequent defining of 1 to N relations on this set. The relations to be defined can reflect (a) inherent structure of the subject matter, (b) pedagogical strategy, (c) successful instructional communication, i.e., the student's current repertoire of subject matter and its structure. It is proposed to represent the set of concepts and relations as graphs or nets, a metalanguage whose mathematical properties are quite well-known. Graph descriptions of instructional subject matter furnish a map so that an instructional agent, human or computer, can orient the presentation. |  |   |

**Security Classification**

| KEY WORDS                          | LINK A |    | LINK B |    | LINK C |    |
|------------------------------------|--------|----|--------|----|--------|----|
|                                    | ROLE   | WT | ROLE   | WT | ROLE   | WT |
| Computer-Administered Instruction  |        |    |        |    |        |    |
| Graph Theory                       |        |    |        |    |        |    |
| Instructional Media and Techniques |        |    |        |    |        |    |
| Metalanguage                       |        |    |        |    |        |    |
| Training Concepts and Theory       |        |    |        |    |        |    |
| Training Systems                   |        |    |        |    |        |    |

Unclassified  
Security Classification